CE88/CP 88: Lecture 1

Data Science for Smart Cities

Design and operation of smart, efficient, and resilient cities nowadays require data science skills.

This course provides an introduction to working with data generated within transportation systems, power grids, communication networks, as well as collected via crowd-sensing and remote sensing technologies.

The core Method of this class Network Science in a framework of Complex Systems thinking to study urban systems. Programming language is Python.

Marta C. Gonzalez

Associate Professor of City and Regional Planning and Civil and Environmental Engineering, Faculty Scientist at the LBNL

My team develops computational models to analyze digital traces to estimate the demand on urban infrastructure in relation to energy and mobility. Examples include: traffic gridlocks and the integration of electric vehicles in the power grid, policy of solar energy adoption, and habits in spending behavior.

http://humnetlab.berkeley.edu/

Alben Rome Bagabaldo

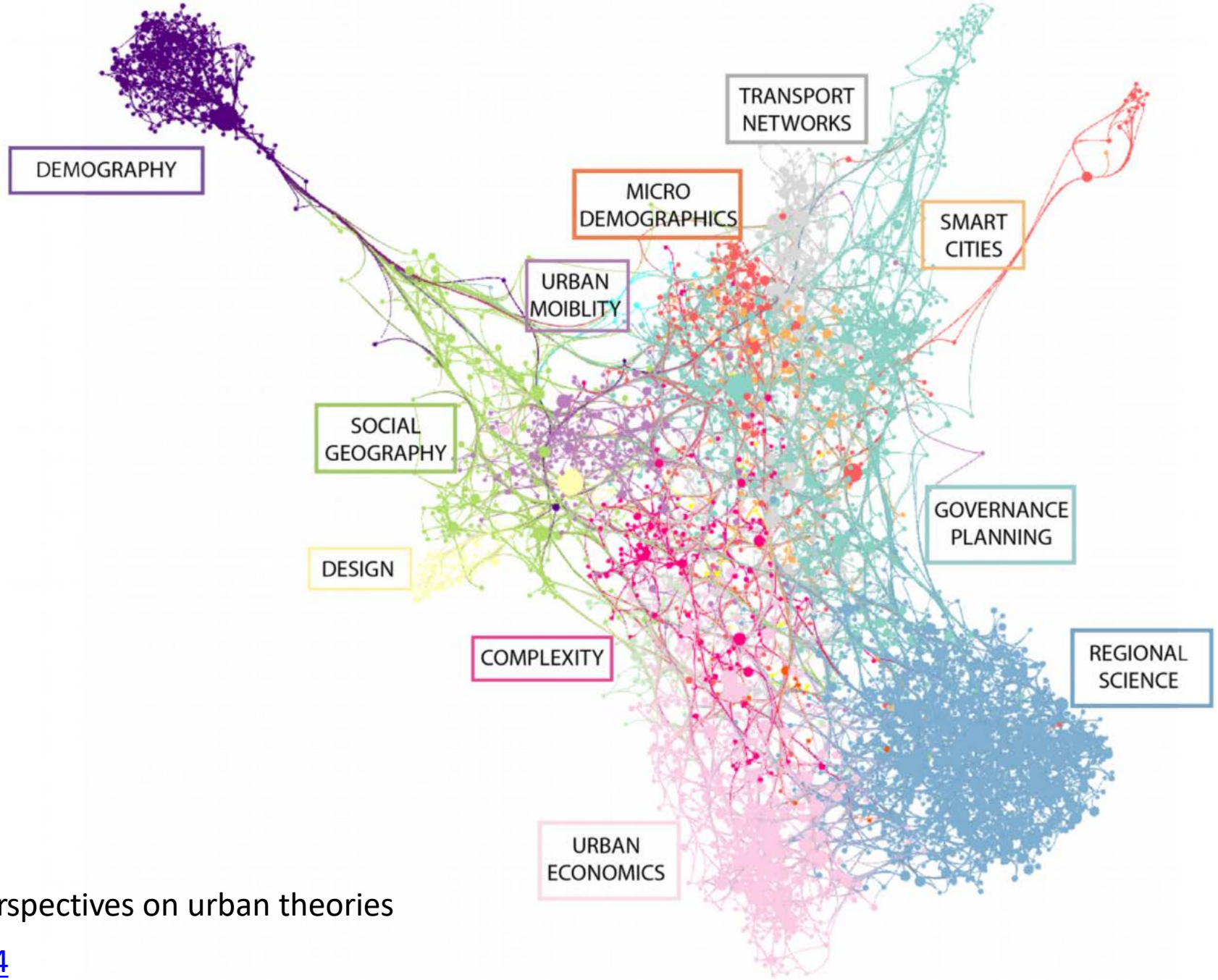
PhD student in Civil and Environmental Engineering (Systems Engineering)
Formerly a full-time faculty member (undergraduate instructor) at the School of Civil, Environmental, & Geological Engineering, Mapua University, Manila Philippines
(https://www.linkedin.com/in/albenbagabaldo/)

Aneesh Diwania

Junior studying Data Science and Statistics with a concentration in Business and Industrial Analytics
(https://aneeshdidwania.com/#projects)

Sarah Zhang

Sophomore majoring in Data Science and Environmental Sciences(https://www.linkedin.com/in/sarahhzhangg/)



Source: Denise Punaim et al., Perspectives on urban theories

https://arxiv.org/abs/1911.02854

CE88: Data Science for Urban Systems

Urban science seeks to understand the fundamental processes that drive, shape and sustain cities and urbanization.

It is a multi/transdisciplinary approach involving concepts, methods and research from the social, natural, engineering and computational sciences, along with the humanities.

source: Report for NSF: Urban Science: Integrated Theory from the First Cities to Sustainable Metropolises

Source: google image

Aerial Image of London

What we find in Cities:

- An Engineered or Natural-physical system
- Social Behavior of people interacting with the system,

and

- Institutional behavio of organized units such as regulators and Markets governing the System.

Note on the report:

Urban analytics is a collection of tools used to analyze and map “urban big data” (generated by social media, crowd sourcing an sensor networks)

Some History

Jane Jacobs as chairperson of a Greenwich Village civic group at a 1961 press conference

Jacobs fought to prevent Washington Square Park, pictured, from being demolished for a highway

The Science of Cities

Jane Jacobs (1916 –2006) journalist, author, and activist who influenced urban studies, sociology, and economics.) arguing that urban renewal did not respect the needs of city-dwellers.

Made us think of cities as problems in organized complexity -- organisms that are replete with unexamined, but obviously intricately interconnected, and surely understandable, relationships.

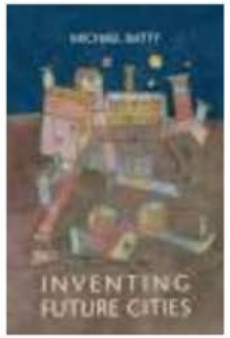
-- **Jane Jacobs *The Life and Death of Great American Cities (1961)***

British Urban Planner

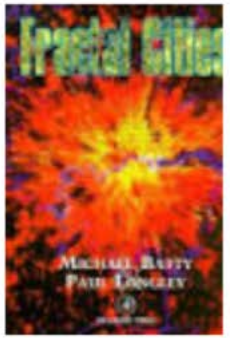
He has pioneered the idea of cities as complex systems as well as for his leading role in the conformation of the Science of Cities that combines a wide spectrum of disciplines ranging from Statistical Physics, Mathematics, Architecture and Engineering, to Social Sciences and Economics.)

The New Science of Cities
2013

Cities and Complexity
2005



Inventing Future Cities
2018

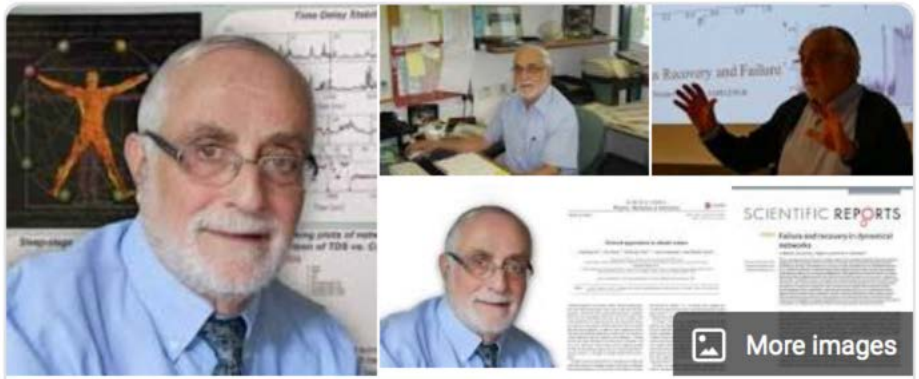


Fractal Cities: A Geometr...
1994

Urban Modelling: Algorith...
1976

Complexity science, also called complex systems science, studies how a large collection of components – locally interacting with each other at small scales – can spontaneously self-organize to exhibit non-trivial global structures and behaviors at larger scales, often without external intervention, central authorities or leaders.

Physics of complex systems connected with Cities…



Shlomo Havlin

Professor

Shlomo Havlin is a Professor in the Department of Physics at Bar-Ilan University, Ramat-Gan, Israel. He served as President of the Israel Physical Society, Dean of Faculty of Exact Sciences, Chairman, Department of Physics. In 2018 he won the Israel Prize for his accomplishments in physics. Wikipedia

H. Eugene Stanley

From Wikipedia, the free encyclopedia

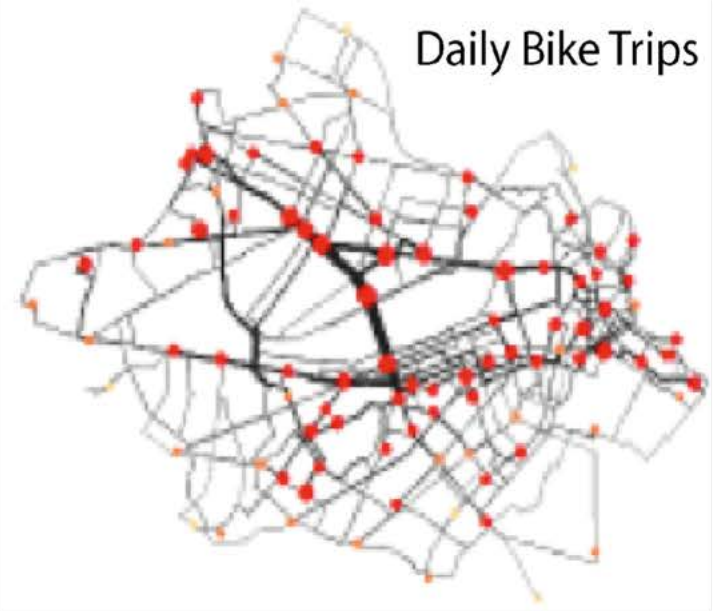
Harry Eugene Stanley (born March 28, 1941) is an American physicist and University Professor at Boston University. He has made seminal contributions to statistical physics and is one of the pioneers of interdisciplinary science. His current research focuses on understanding the anomalous behavior of liquid water, but he had made fundamental contributions to complex systems, such as quantifying correlations among the constituents of the Alzheimer brain, and quantifying fluctuations in noncoding and coding DNA sequences, interbeat intervals of the healthy and diseased heart. He is one of the founding fathers of econophysics.

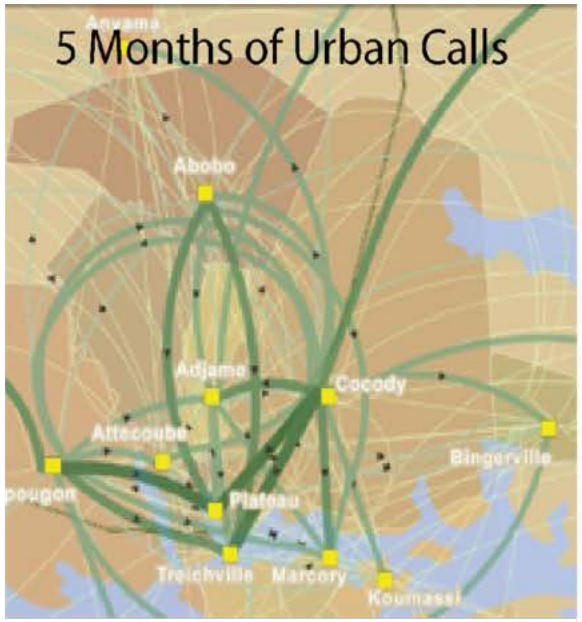
Interactive Examples: https://complexityexplained.github.io/

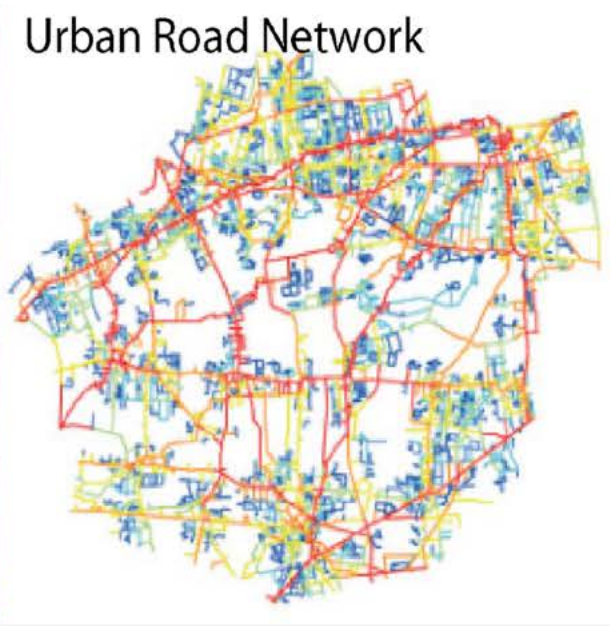
Examples: billions of interacting neurons in the human brain; computers communicating in the Internet; humans in multifaceted relationships.

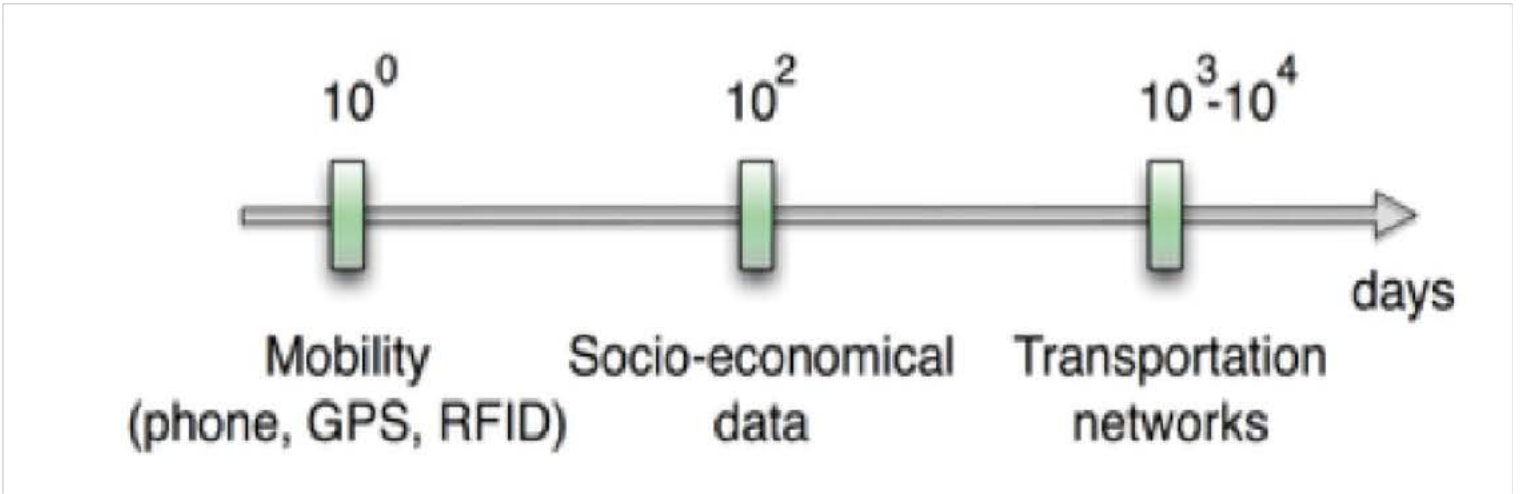
Concepts: system, component, interactions, network, structure, heterogeneity, inter-relatedness, inter-connectedness, interdependence, subsystems, boundaries, environment, open/closed systems, systems of systems

Booklet: https://**complexityexplained**.github.io/ComplexityExplained.pdf



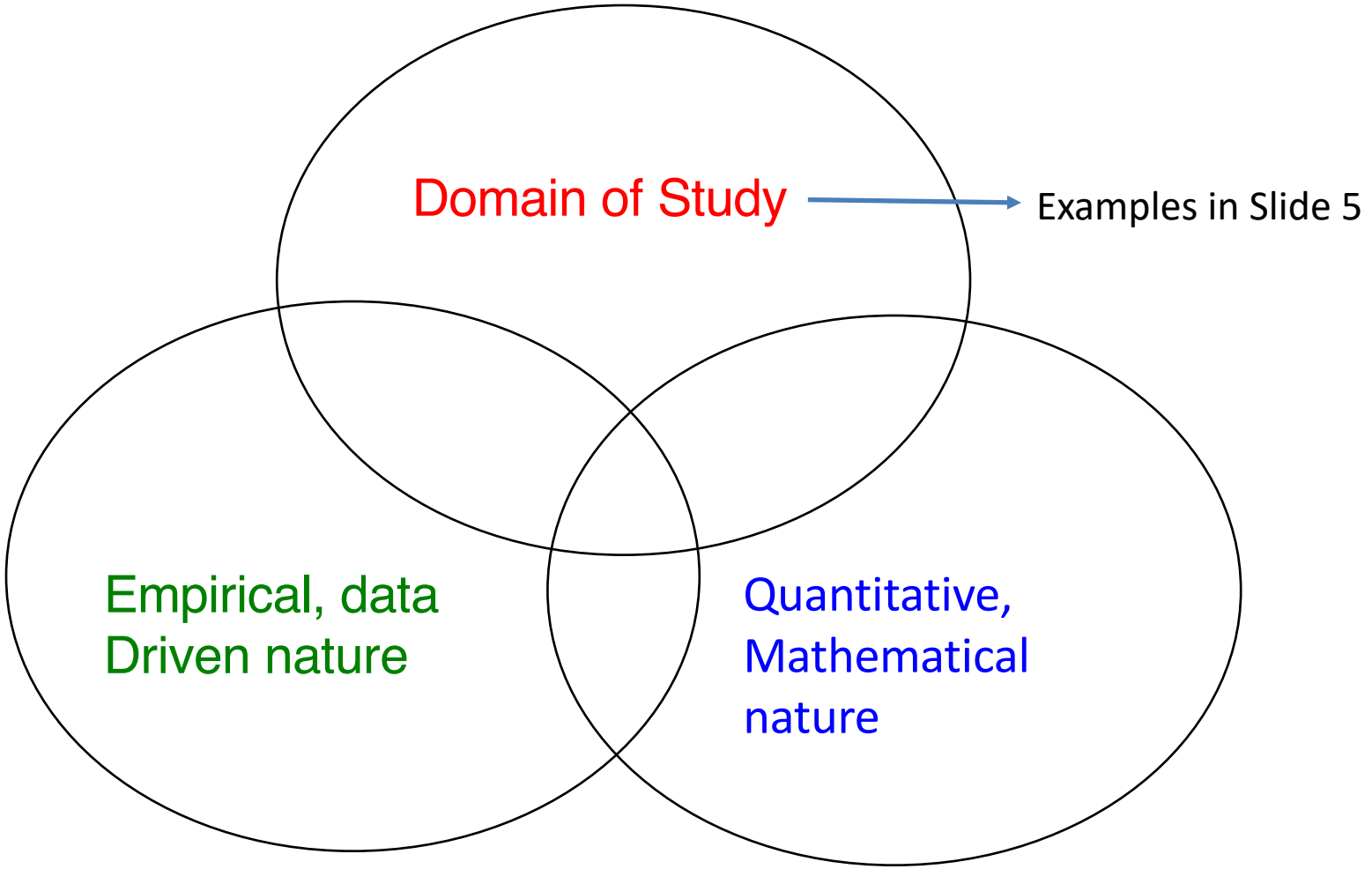






Review "Spatial Networks" published in Physics Reports (2011) Free online: https://arxivorg/abs/1010.0302
Book: "Morphogenesis of Spatial Networks" Springer (2018)

Characteristics of Network Science



The core Method of CE88: Network Science in a framework of Complex Systems thinking to study urban systems.

Weeks 1-2: Network and Small Worlds
(Assignment 1: Social Networks)
Weeks 3-4: Network Centralities,
OpenStreet Maps in NetworkX
(Assignment 2: Subway Networks)
Weeks 5-6: Modularity and Clusters in Networks
(Assignment 3: Migration Networks)
Week 7-8: L7 Spatial Analysis of Census Data and K-means Clustering
(Assignment 4: Spatial Analysis of Census Data)
Week 9: Data Visualization
Week 10: Spreading Dynamics
(Assignment 5: Project Preparation and Paper Drafts)
Week 11: Class Review
(Final Presentation (5%) and Project Paper (10%))

Assignments 1-4: 60%
Participation: 10%
Project Preparation: 15%
Final Paper: 10% and Presentation 5%

Formal definitions

- A **network** or **graph** *G* has two parts, a set of *N* elements, called **nodes** or **vertices**, and a set of *L* pairs of nodes, called **links** or **edges**. The link *(i, j)* joins the nodes *i* and *j*. Two nodes are **adjacent** or **connected** or **neighbors** if there is a link between them.

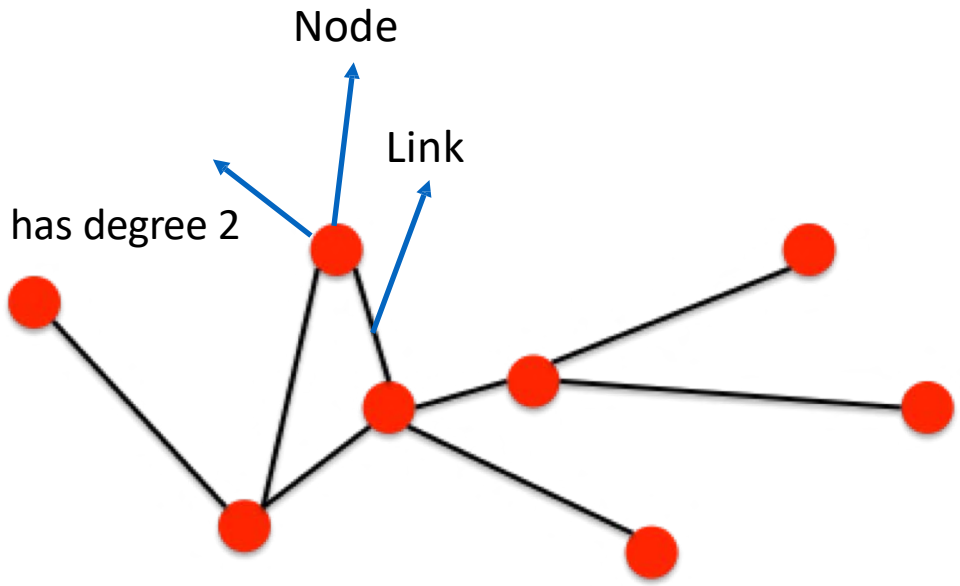
- A network can be **undirected** or **directed**. A directed network is also called a **digraph**. In directed networks, links are called **directed links** and the order of the nodes in a link reflects the direction: the link *(i, j)* goes from the **source** node *i* to the **target** node *j*. In undirected networks, all links are bi-directional and the order of the two nodes in a link does not matter.

- A network can be **unweighted** or **weighted**. In a weighted network, links have associated **weights**: the **weighted link** *(i,j,w)* between nodes *i* and *j* has weight *w*. A network can be both directed and weighted, in which case it has directed weighted links.

Source: A first course in Network Science Chapter 1

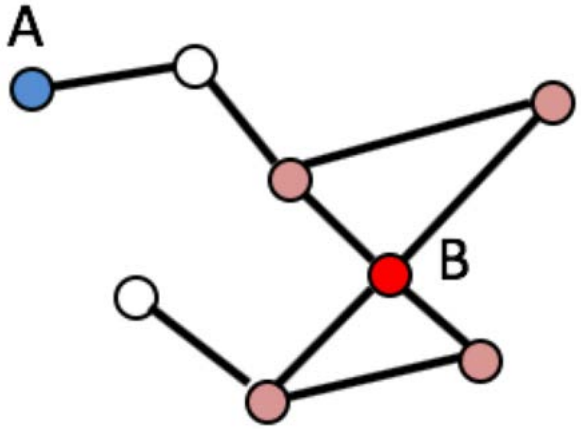
In Summary:

- A **network** (graph) is a structure used to model pairwise relations between objects. A network is made up of entities (aka **nodes**, vertices) and the relationships between them (aka **links**, edges)
- Number of edges of incident to the vertex is **the node degree**

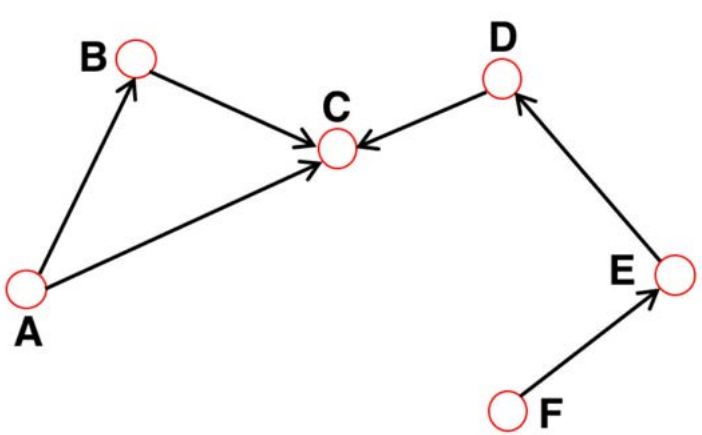


Network Types

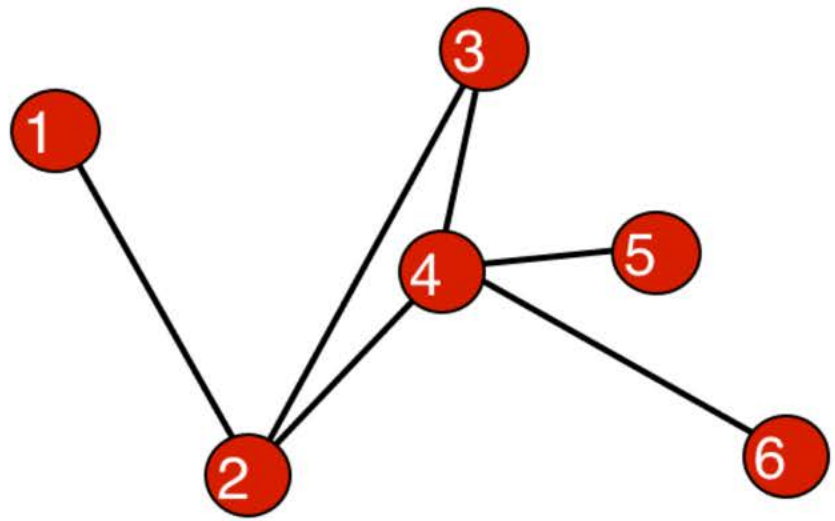
- Undirected



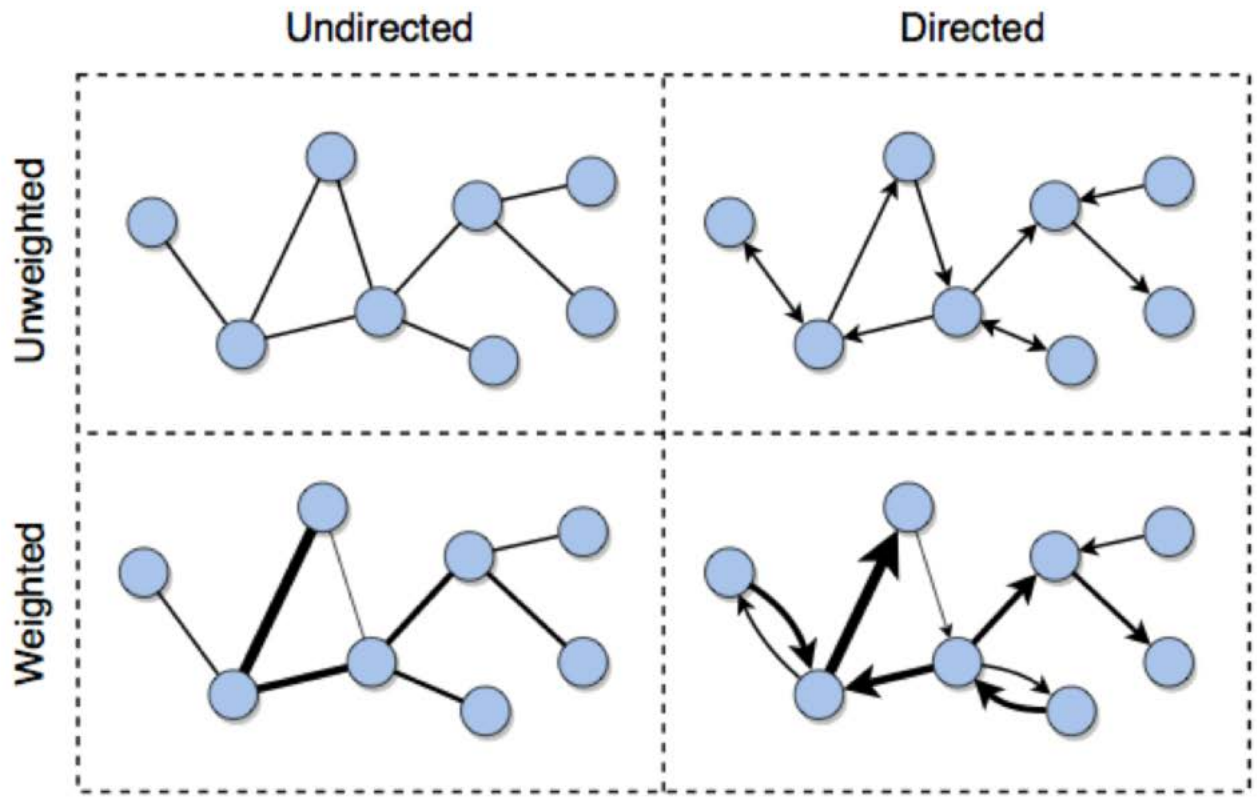
- Directed



Python and NetworkX

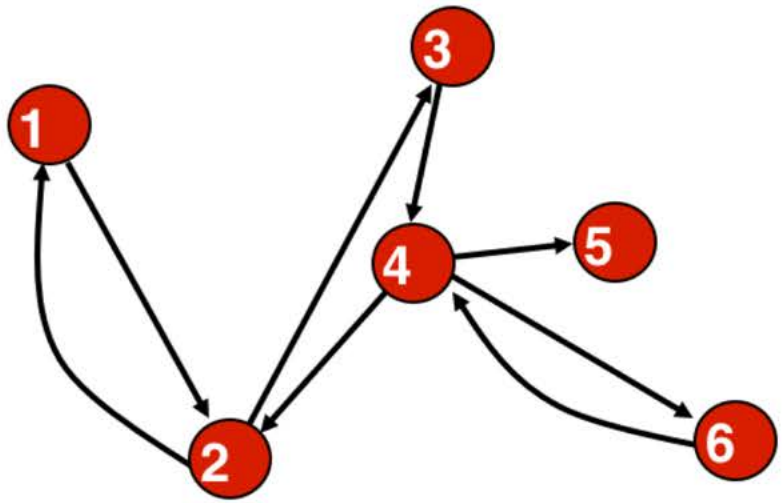



```
import networkx as nx # always!

G = nx.Graph()
G.add_node(1)
G.add_nodes_from([2,3,…])
…
G.add_edge(1,2)
G.add_edges_from([(2,3),(2,4),…])
…
G.nodes()
G.edges()
G.neighbors(4)

for n in G.nodes:
    print(n, G.neighbors(n))
for u,v in G.edges:
    print(u, v)
```



Can you think of a few examples in each of these categories?

Directed networks




```
import networkx as nx # don't forget!

D = nx.DiGraph()
D.add_edge(1,2)
D.add_edge(2,1)
D.add_edges_from([(2,3),(3,4),…])
…
D.number_of_nodes()
D.number_of_edges()
D.edges()
D.successors(2)
D.predecessors(2)
D.neighbors(2)
```

Activity 1:

Separate in 12 Groups and discuss the Networks in slides 22 to 33 reply the posted questions.

Room1


Baran
Abali
3033549822


Ergun
Acikoz
3036297874


Shreya
Ayyagari
3032673100


Dea
Bardhoshi
3034820868


Christal
Bermudez
3032801371


Erin
Bhan
3035815574

Eugene
Brodsky
3033401516


Nadav
Bronicki
3034520217


Joseph
Casey
3034390935


Austin
Chang
3032706614


Amanda
Chen
3032717651


Katerine
Chung Chen
3034206881


Jonathan
Dena
3036185177


Cosmin
Deshmukh
3034551612


Maya
Facciolo
3034248884


Karina
French
3035305651


JP
Garcia
3036324121


Daniel
Ha
3033710736


Valerie
Haines
3033707148


Drake
Hayes
3032825980

Vanessa
Hernandez-Deniz
3035105178

Cheng-Kai
Hsu
3036245367


Kaia
Hu
3034529551


Kaichi
Inoue
3036002943


Isabella
Isles
3033732199


Zhiyong
Jiang
3035418598


Daniel
Kennedy
21656790


Alice
Lara
3035705568


Ying Chik
Lee
3035350202


Amelia
Li
3035149144

Room6


Angela
Li


Julian
Lichtenfeld


Jodie
Lu


Varsha
Madapoosi


Devon
Madeksiak


Anaka
Maher

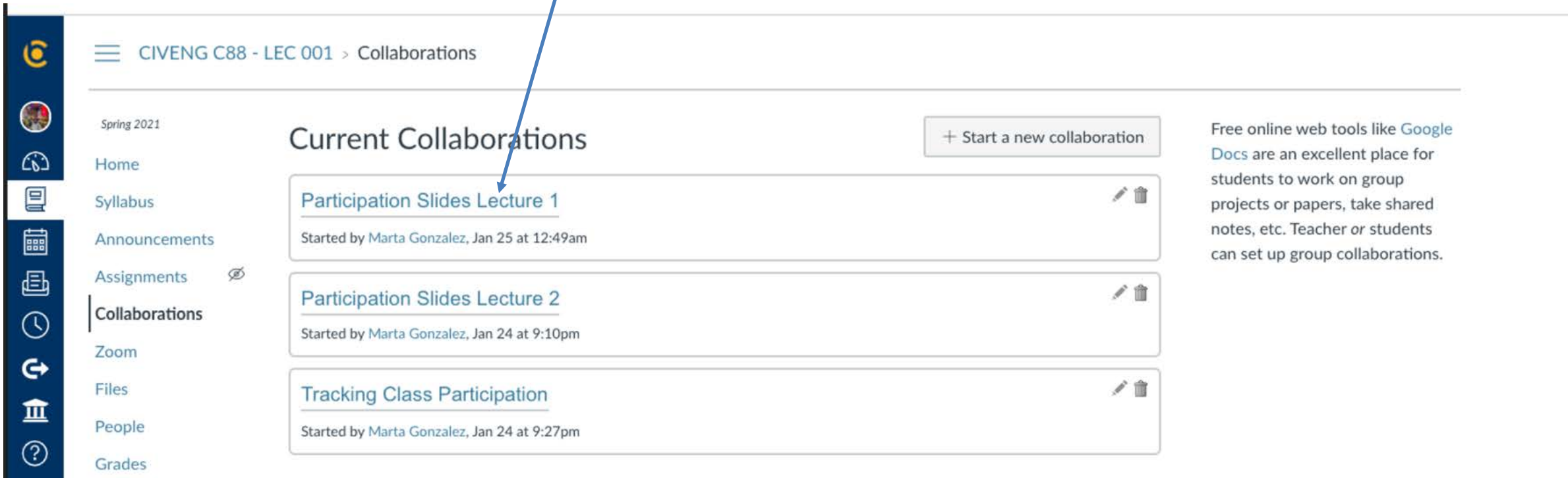
Room 7

3032712243	3034847622	3032634789	3034759768	3033950027	3036318648
					No Official Photo is Available
Pranav **Mallampalli** 3033954954	Emmiee **Malyugina** 3032662492	Sydney **Maves** 3036267987	Rachel **McCarty** 3034642196	Leah **Mealey** 3034147536	Faizaan **Merchant** 3036185749
			No Official Photo is Available	No Official Photo is Available	No Official Photo is Available
Juan **Montoya Hernandez** 3032721343	Tasnima **Naoshin** 3033840879	Rohan **Nath** 3032919879	Aaron **Opell** 3033837928	Gerald **Powell** 25009424	Felicia **Pramudji** 3035033912
		No Official Photo is Available			
Emilio **Recinos-Walsh** 3035096000	Sahil **Sanghvi** 3032638533	Maggie **Siu** 3034195506	Sena **Soysal** 3035252039	Philippa **Steinberg** 3034932590	Stephany **Su** 3033977184
			No Official Photo is Available	No Official Photo is Available	
Daniel **Sun** 3032643135	Jovin **Thomas** 3035833358	Paloma **Torres** 3035612358	Justin **Wong** 3032698580	Lucius **Wu** 3036189311	Sven **Wu** 3032645059
					
Brian **Xi** 3035923903	Timothy **Yang** 3034265459	Kyaw Swar **Ye Myint** 3036214921	Jarvis **Yuan** 3034041365	Brandon **Yung** 3034022593	Maxwell **Zinkievich** 3034071837

Room 11

If you are not in the Roster yet select Room 12

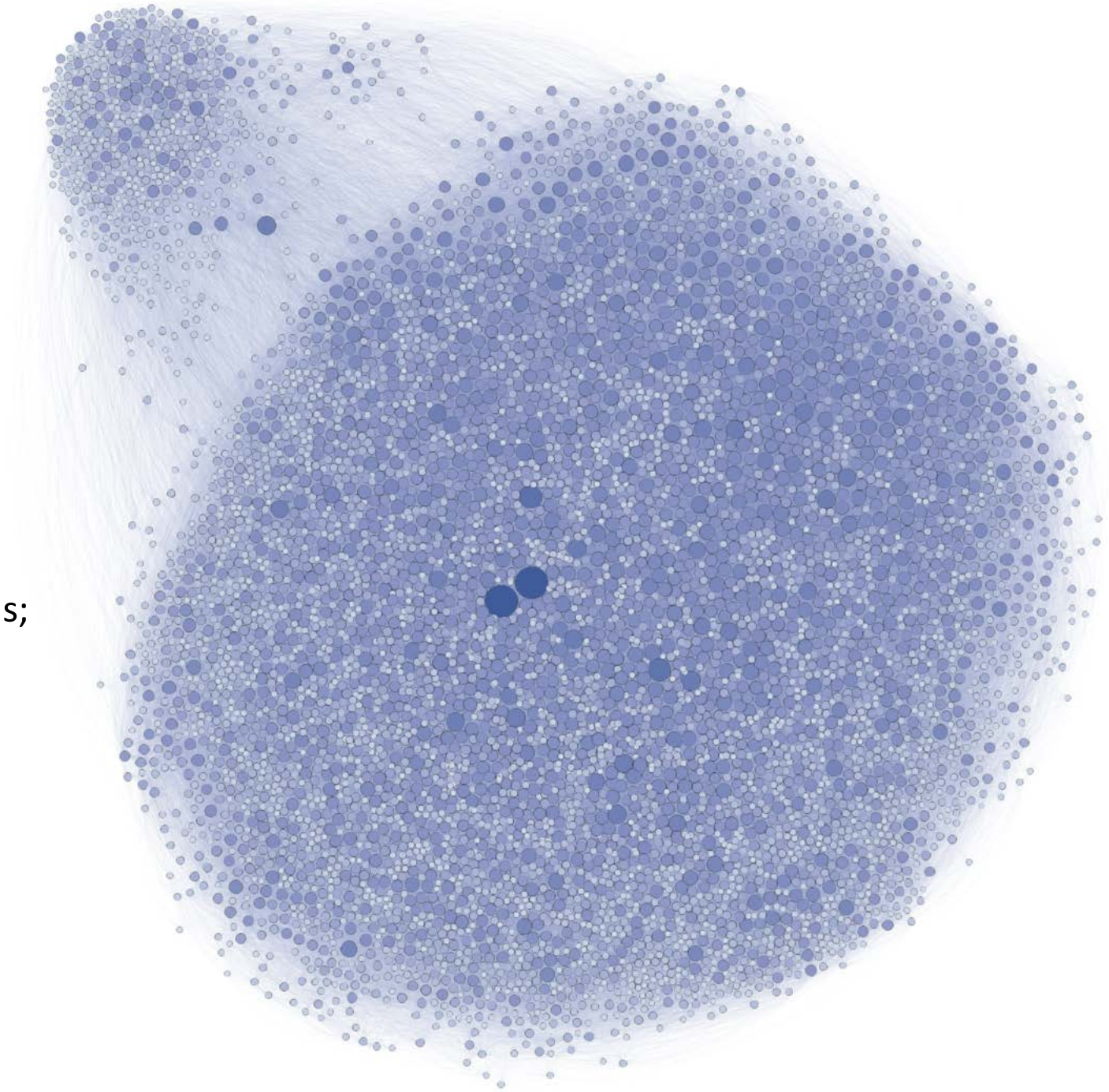
Click there to open the slides and add the answers

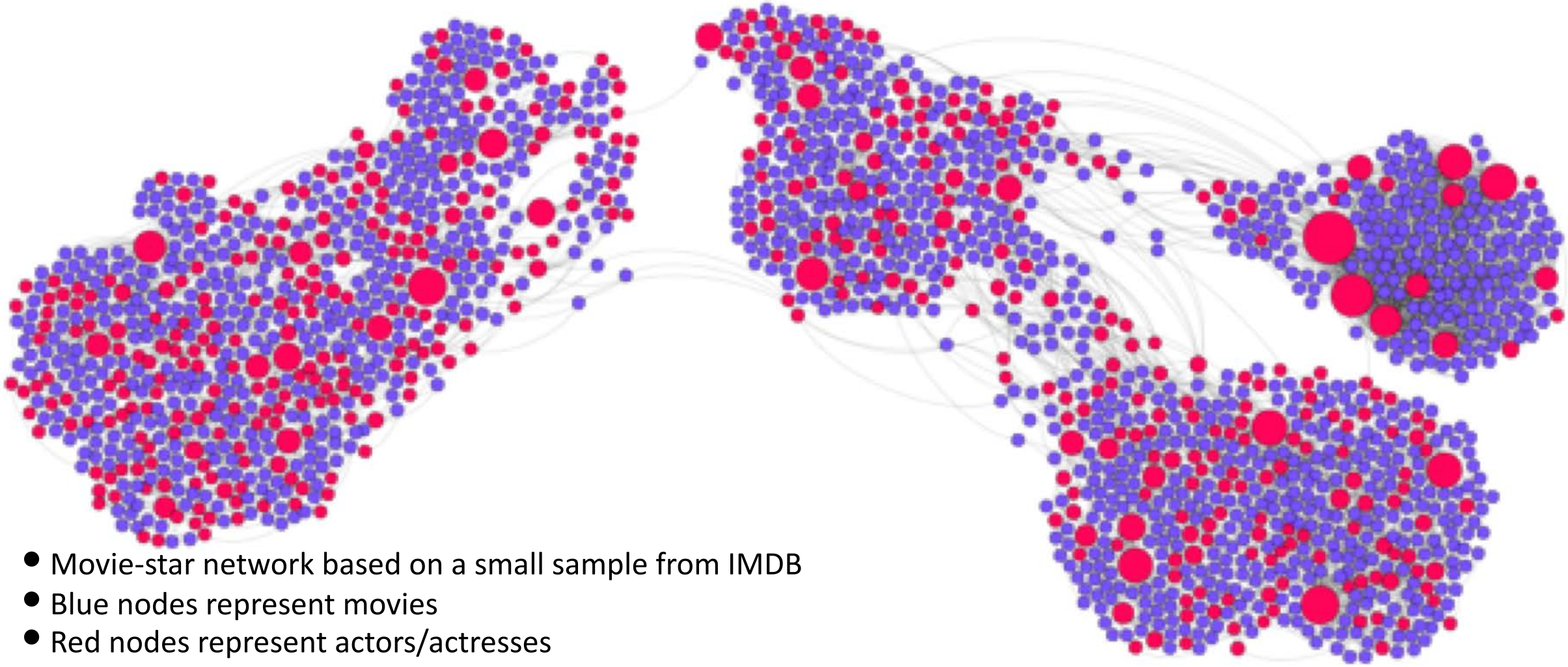


https://docs.google.com/presentation/d/1GDjVUUvgFOBS9QfG28E6Y_DVJuuETFeWtB1tBsSlHsI/edit?usp=sharing

- Facebook users at Northwestern University
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger, darker nodes have more connections; what does that represent?
- What do the two clusters tell us?

Bonus: Approx. number of nodes and Links?



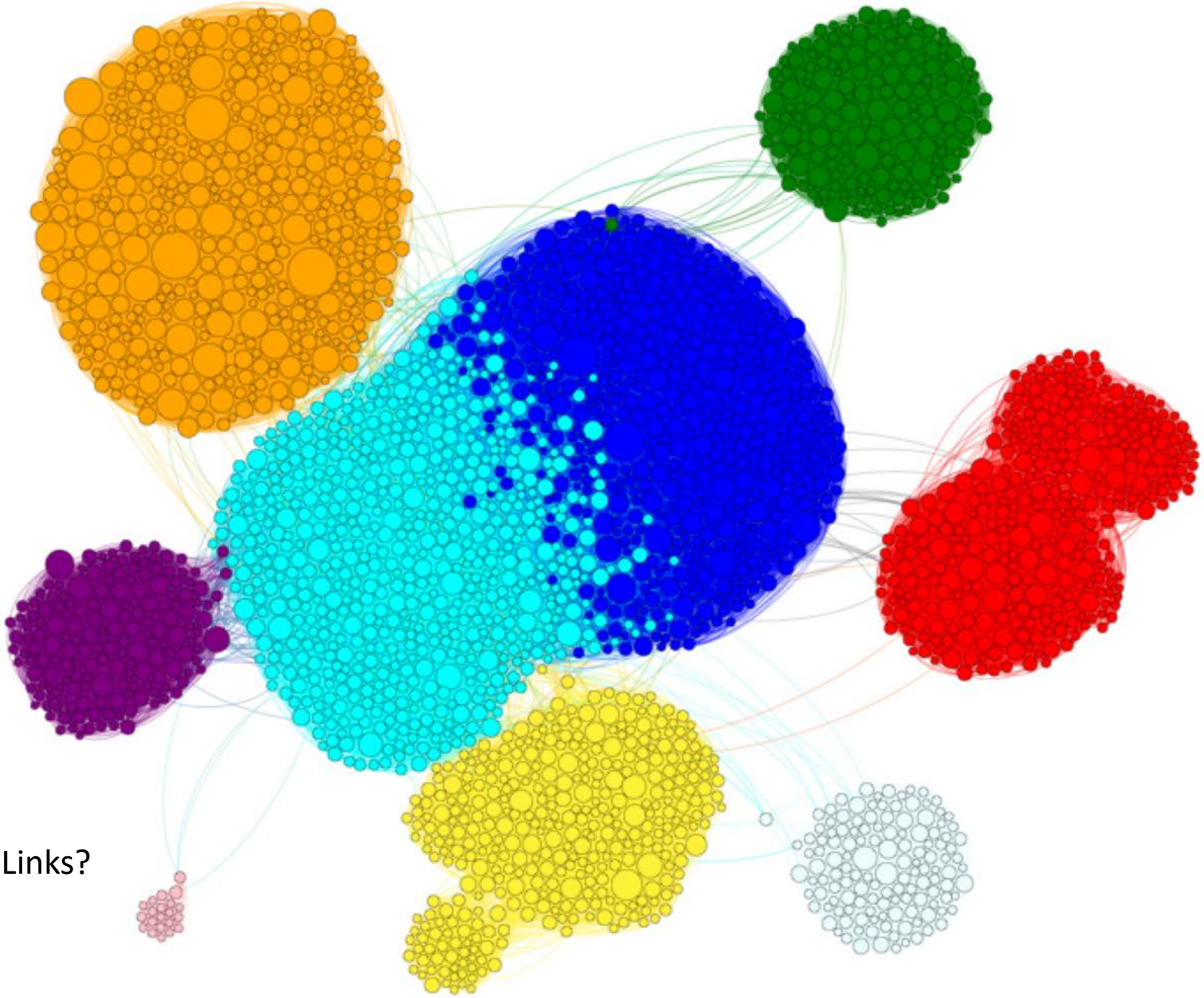


- Movie-star network based on a small sample from IMDB
- Blue nodes represent movies
- Red nodes represent actors/actresses
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes have more connections; what does that mean?
- What do the clusters represent?

Bonus: Approx. number of nodes and Links?

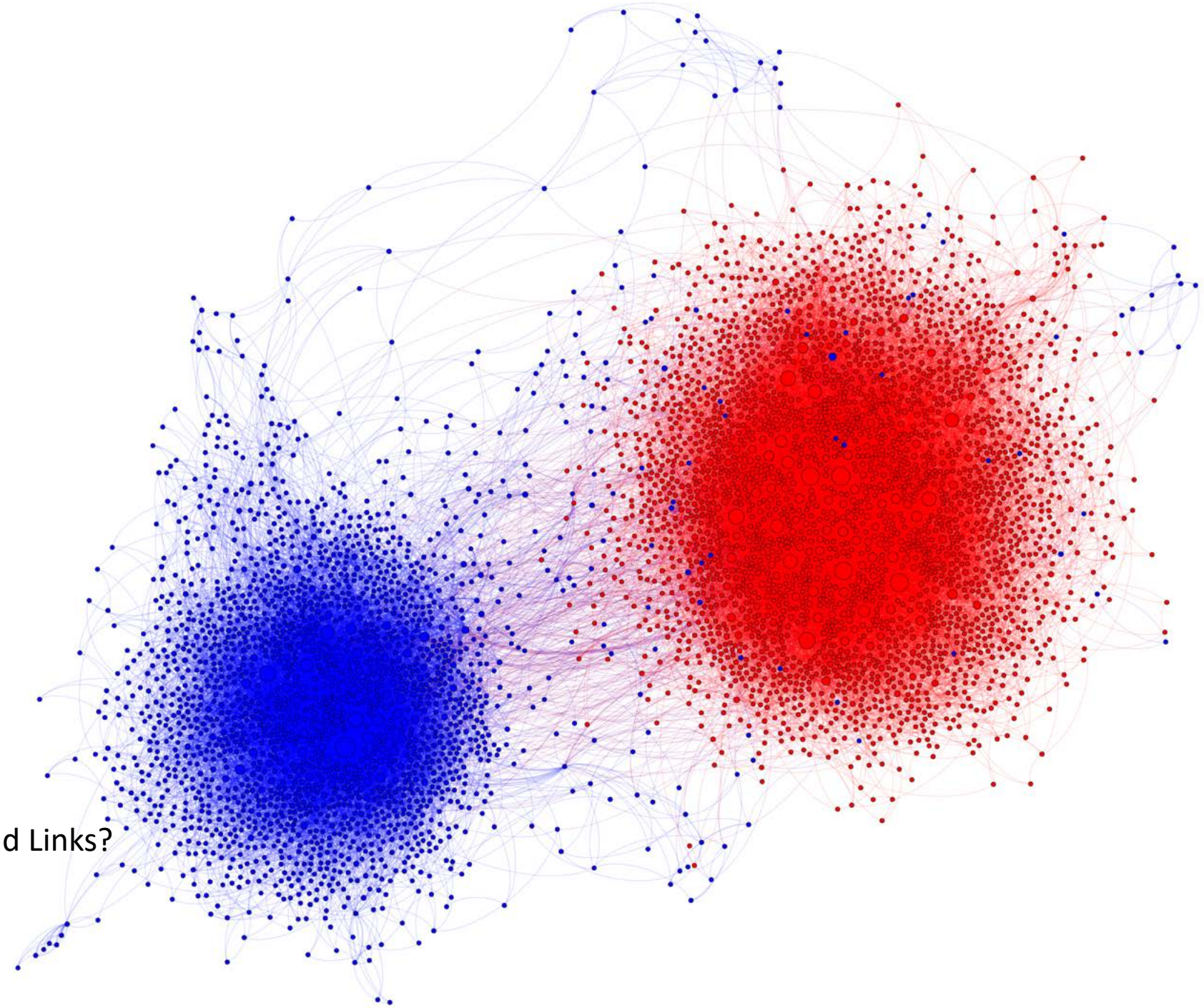
- Movie co-star network based on a small sample from IMDB
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes have more connections; what does that mean?
- What do the clusters represent?

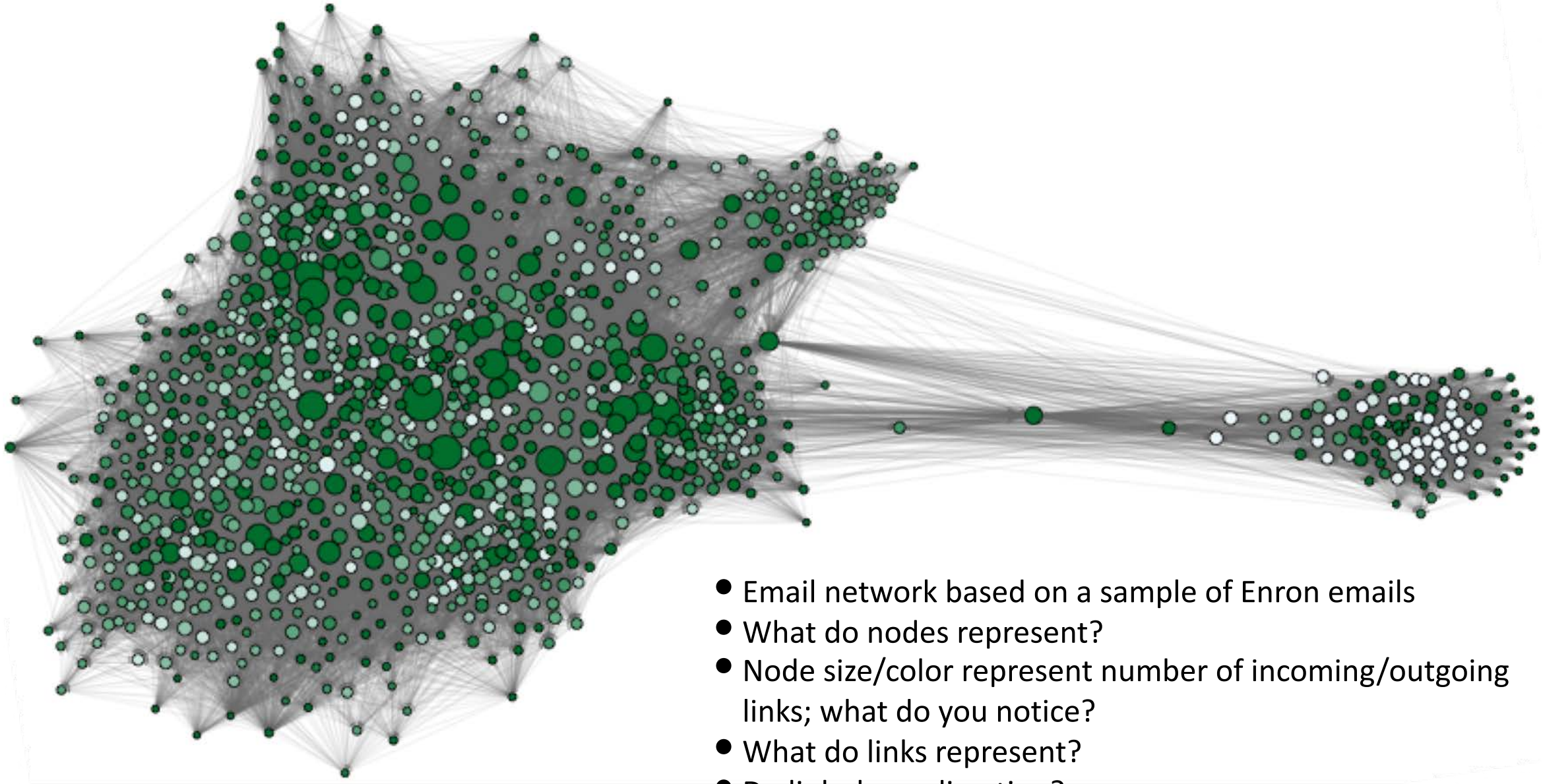
Bonus: Approx. number of nodes and Links?



- Retweet network on Twitter, based on political posts during 2010 US election
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes have more connections; what does that mean?
- What do the clusters and colors represent?

Bonus: Approx. number of nodes and Links?

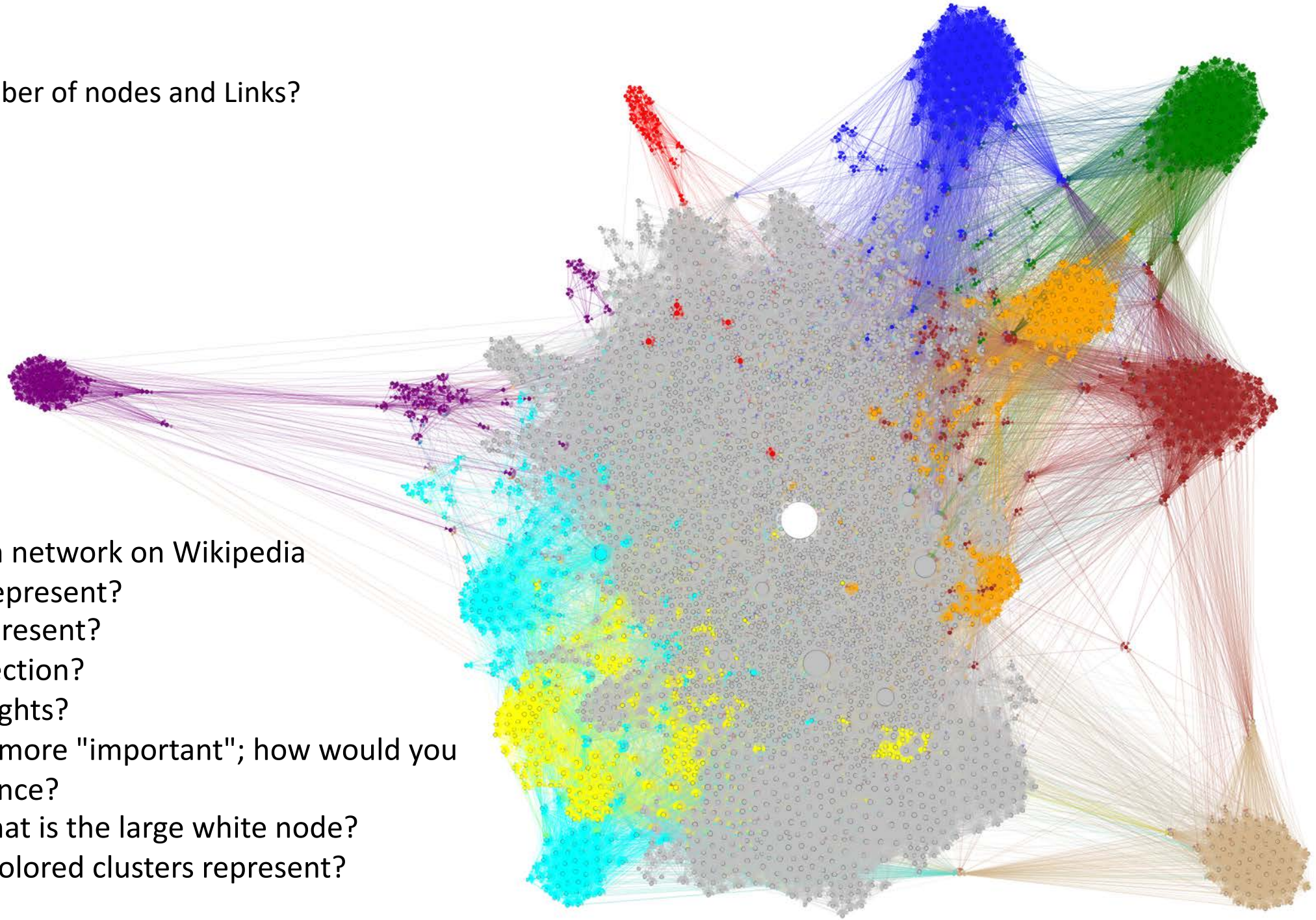




- Email network based on a sample of Enron emails
- What do nodes represent?
- Node size/color represent number of incoming/outgoing links; what do you notice?
- What do links represent?
- Do links have direction?
- Do links have weights?
- What do the clusters represent?

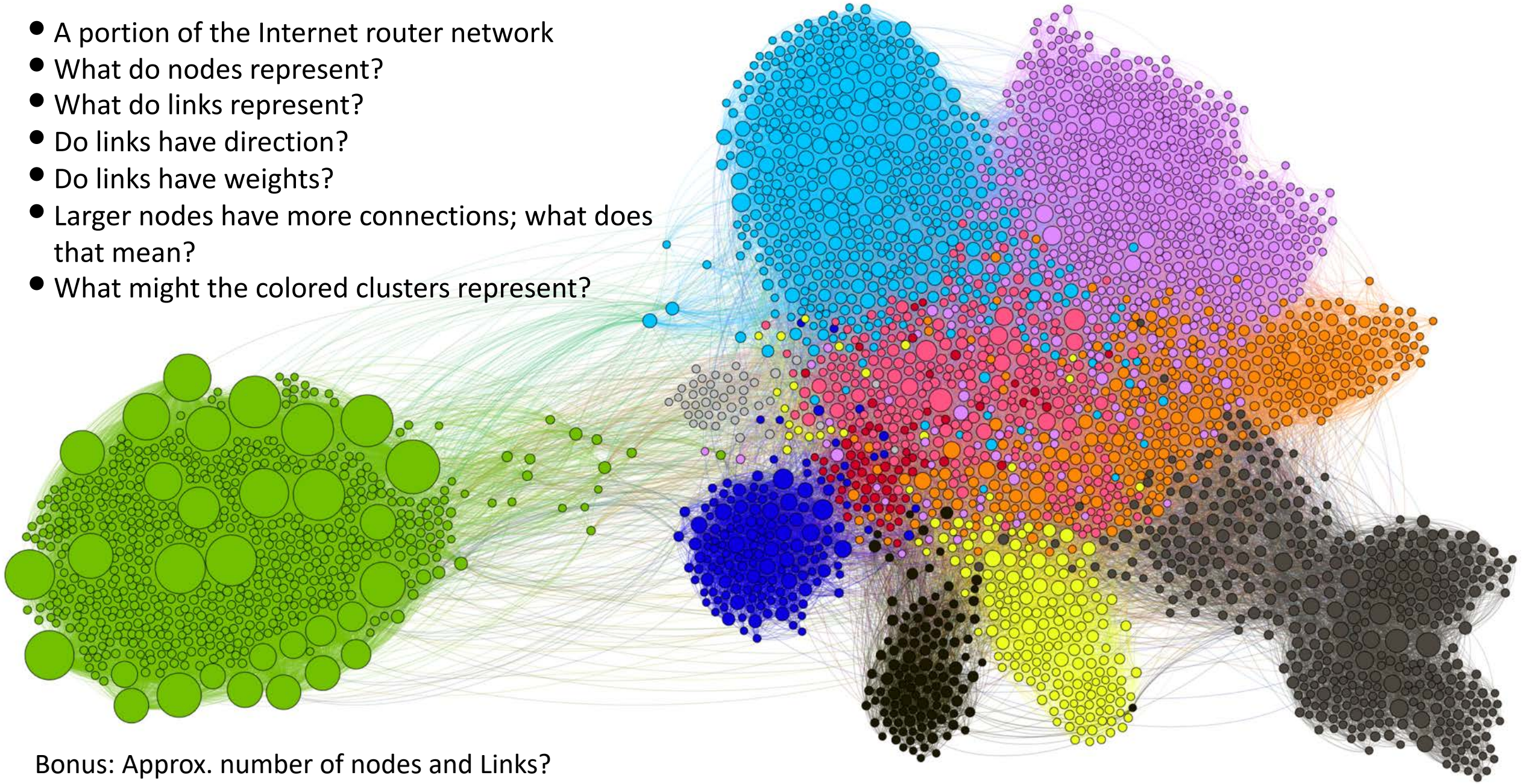
Bonus: Approx. number of nodes and Links?

Bonus: Approx. number of nodes and Links?

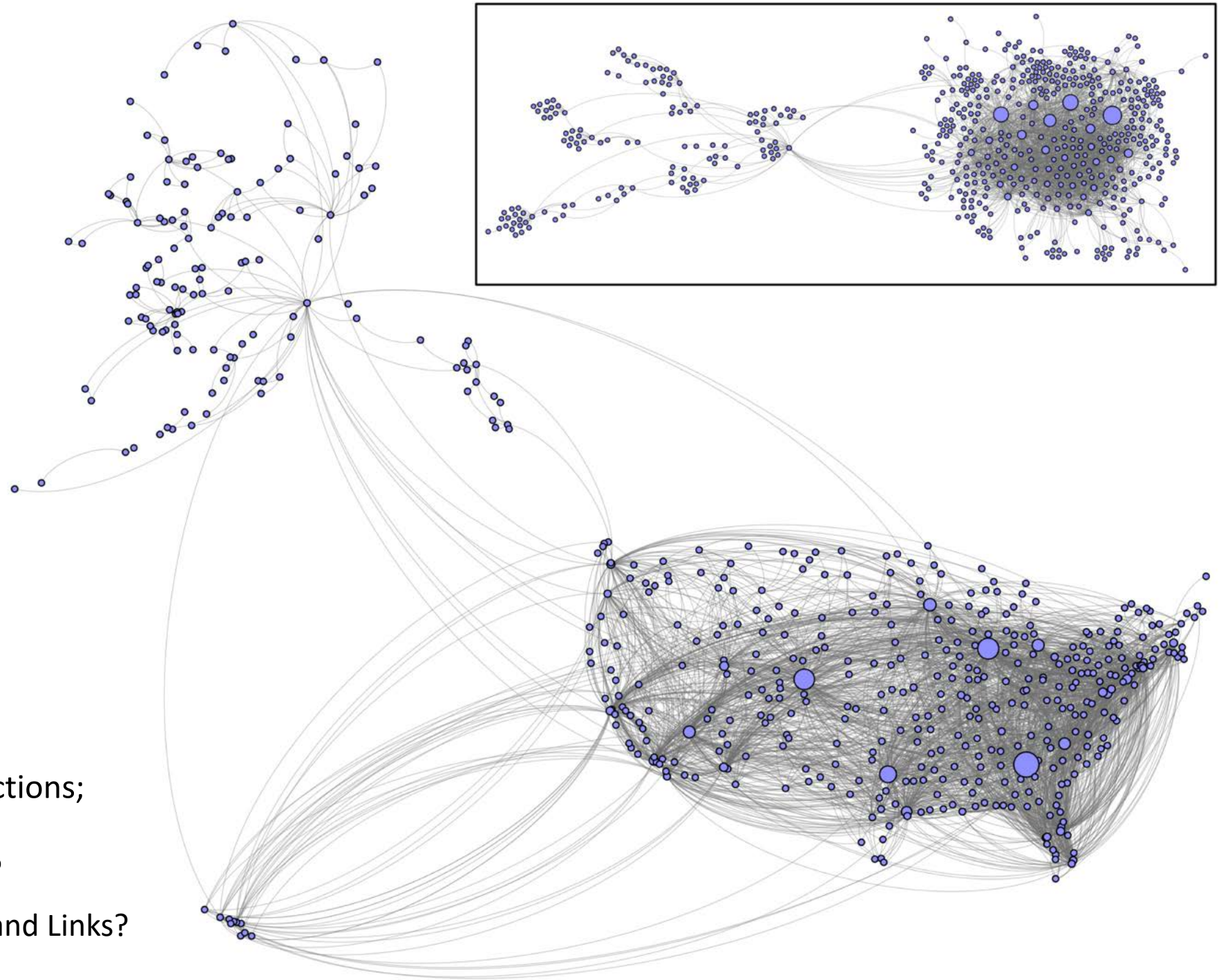


- Math information network on Wikipedia
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes are more "important"; how would you measure importance?
- Can you guess what is the large white node?
- What might the colored clusters represent?

- A portion of the Internet router network
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes have more connections; what does that mean?
- What might the colored clusters represent?

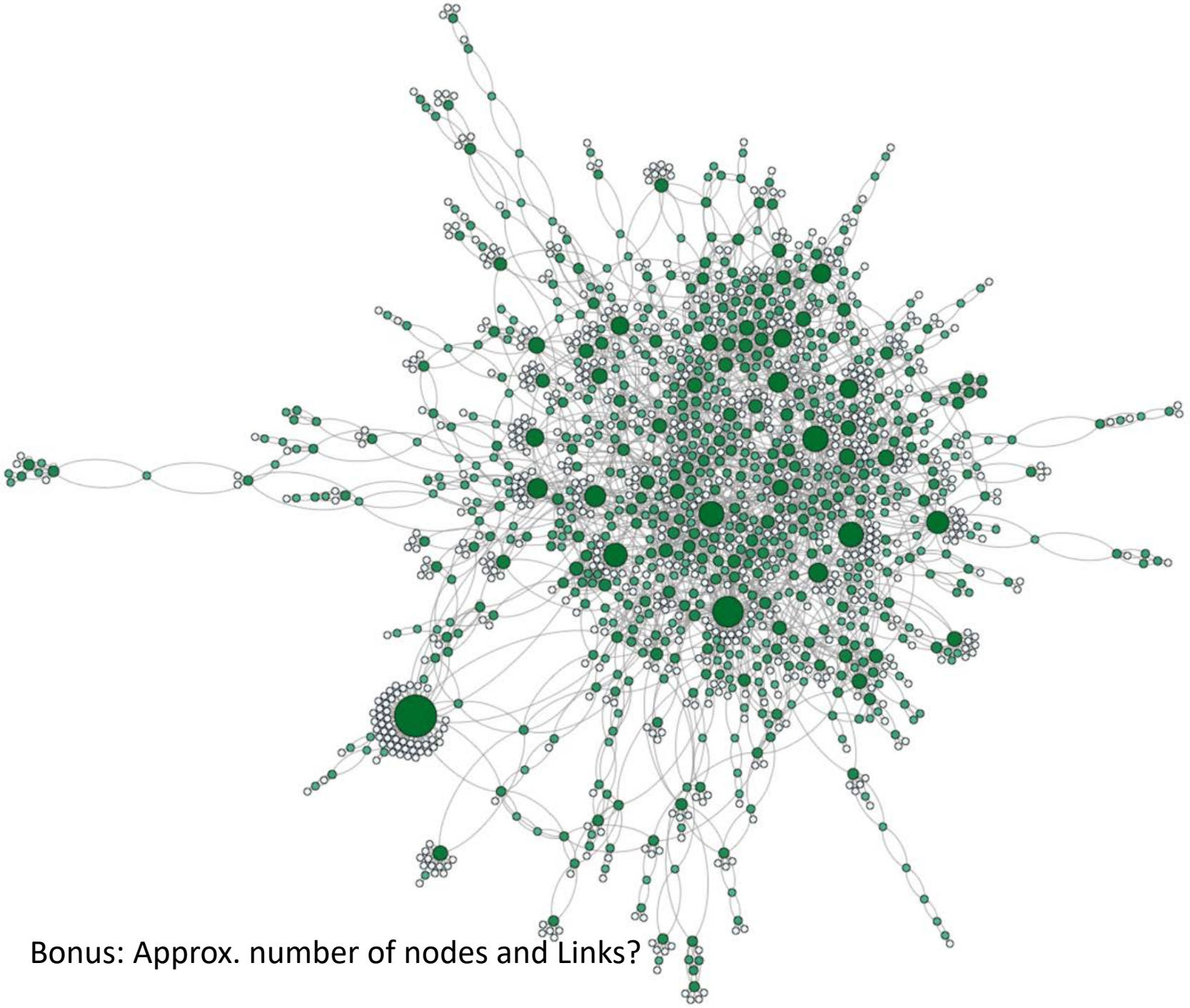


Bonus: Approx. number of nodes and Links?



- US air transportation network
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes have more connections; what do they represent?
- What do the layouts represent?

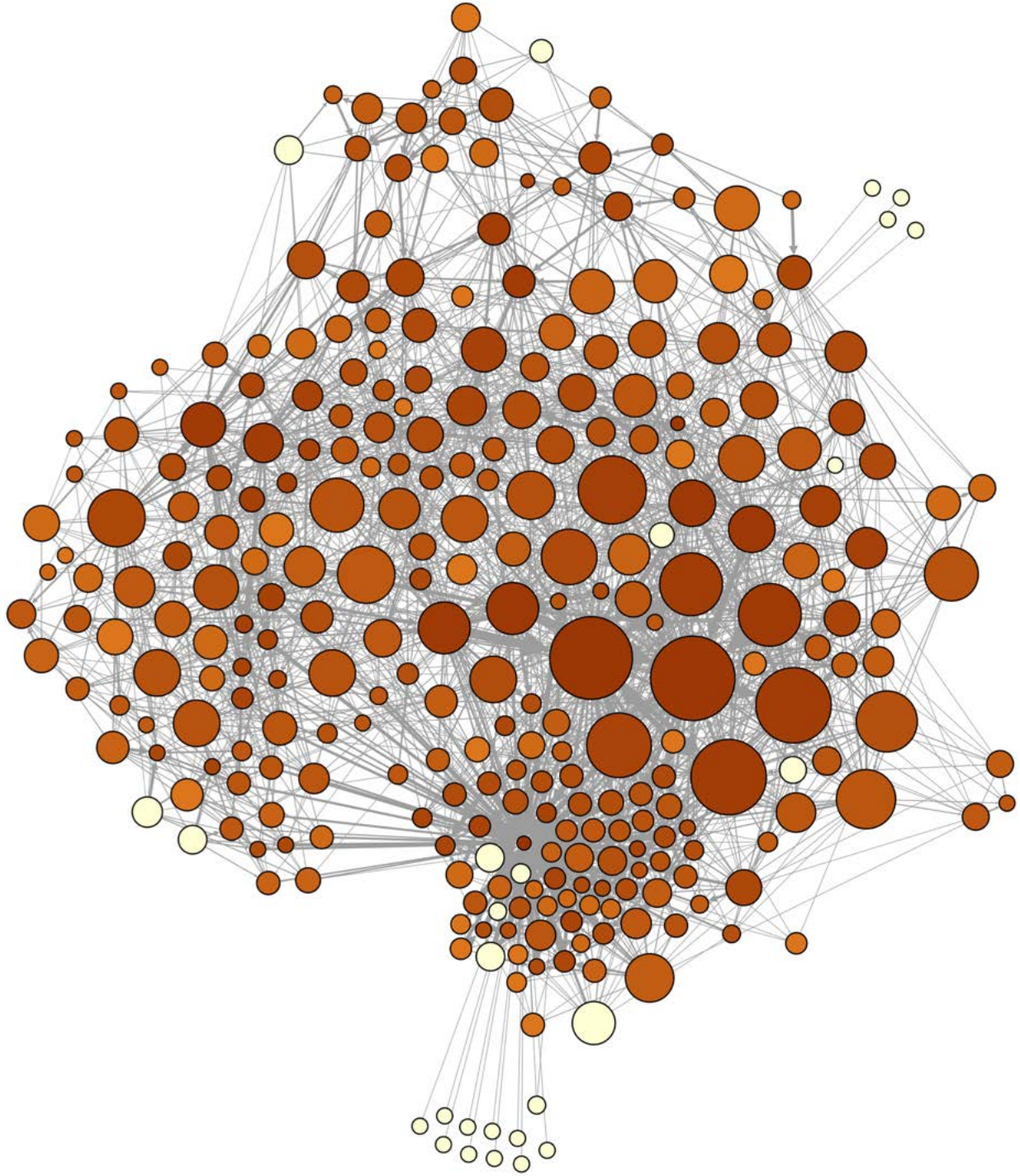
Bonus: Approx. number of nodes and Links?



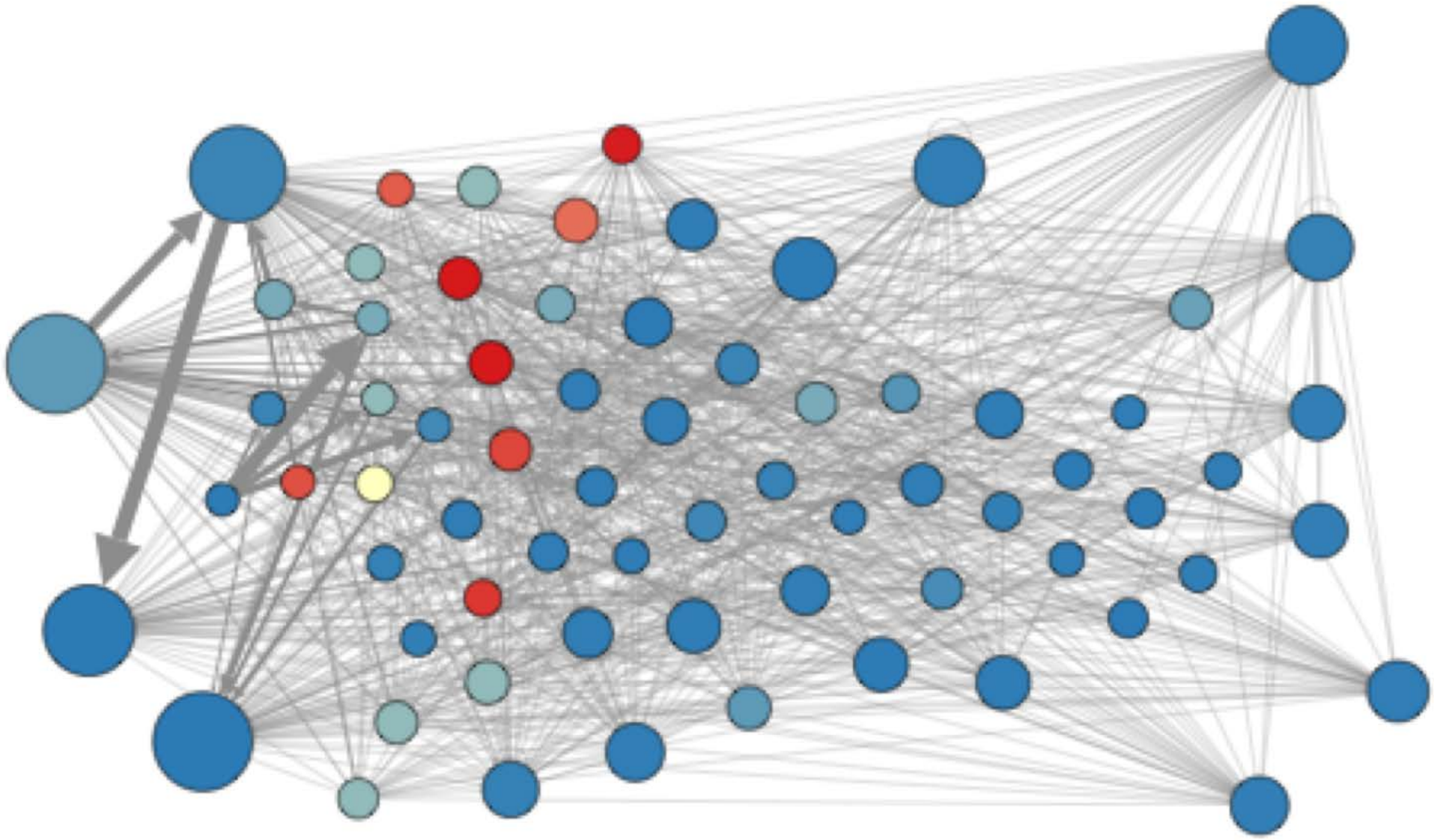
- Protein interaction network of yeast
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger nodes have more connections; what does that mean?
- What do the clusters represent?

Bonus: Approx. number of nodes and Links?

- Neural network of the roundworm c. elegans
- What do nodes represent?
- What do links represent?
- Do links have direction?
- Do links have weights?
- Larger/darker nodes have more outgoing/incoming connections; what does that mean?



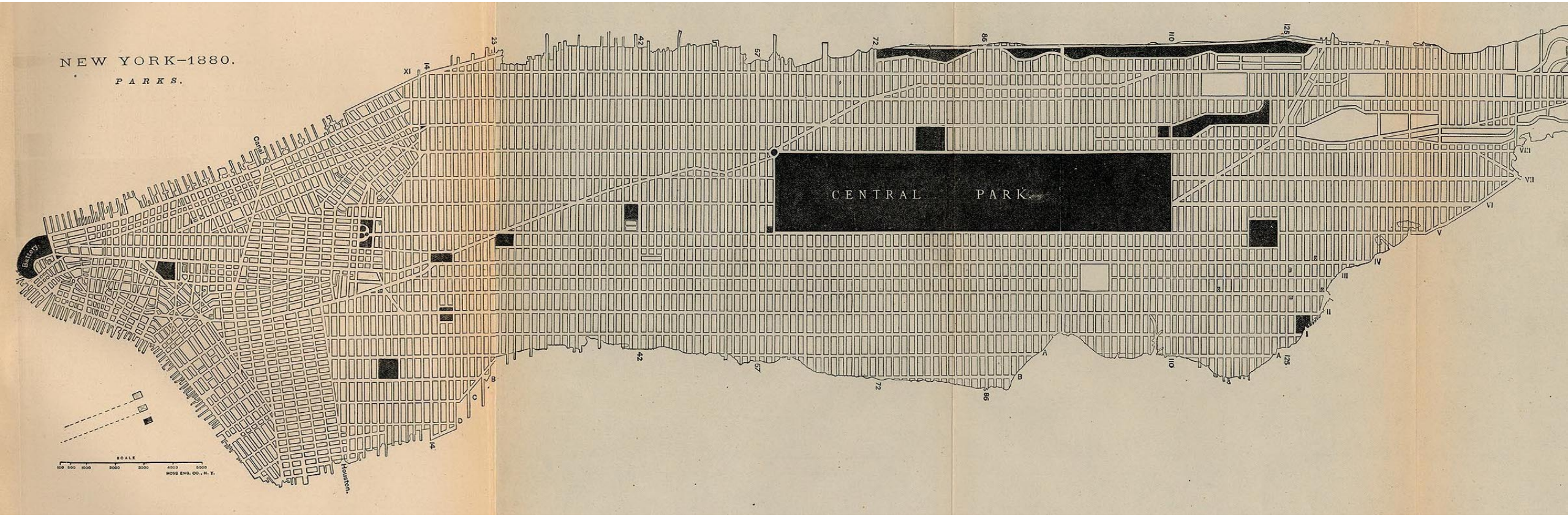
Bonus: Approx. number of nodes and Links?



- Food web of species in the Florida Everglades
- What do nodes represent?
- What do links represent?
- Do links have direction? What does it represent?
- Do links have weights? What do they represent?
- Larger nodes have more incoming links; what are they?
- Red nodes have more outgoing links; what are they?

Bonus: Approx. number of nodes and Links?

What are the nodes and links of the network represented in this street map?



Map of New York in 1880. From Report on the Social Statistics of Cities, Compiled by George E. Waring, Jr., United States Census Office, 1886. Image courtesy of University of Texas Libraries

Bonus: Approx. number of nodes and Links?

Activity 2:
See the Installation Instructions uploaded in bcourses and open the first Script

MyFirstNetwork_Exercise-woSoln.ipynb

Activity 3:
Using the documentation of NetworkX here:

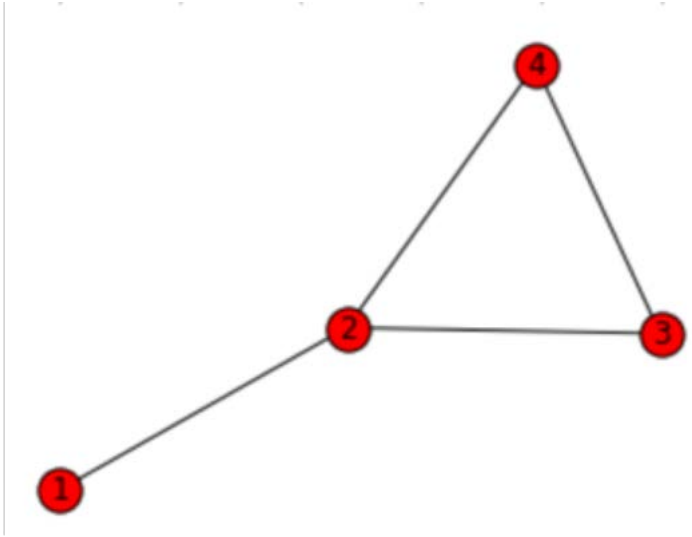
https://networkx.org/documentation/stable/tutorial.html#nodes

Fill the answers of MyFirstNetwork_Exercise-woSoln.ipynb

(you can read MyFirstNetwork_Exercise-woSoln.html while you complete installations after class)

Undircted Graph

1) Nodes
2) Edges
3) Number of Nodes
4) Number of Edges
5) Number of Neighbors per node
6) Degree of each node
7) Average degree calculation and short equation



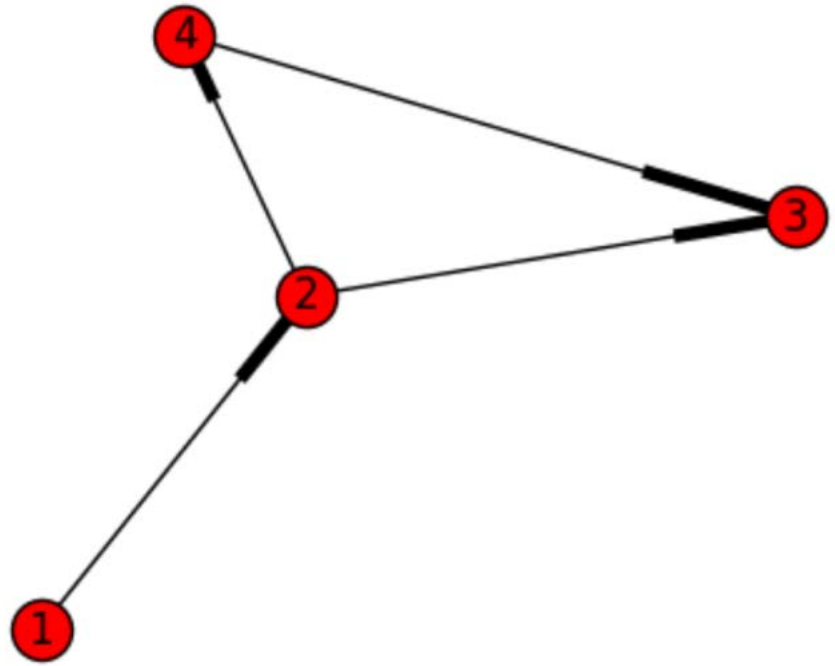
Directed Graph

1) Nodes
2) Edges
3) Number of Nodes
4) Number of Edges
5) Number of Neighbors per node
6) Degree of each node

   In-degree of each node

   out-degree of each node

7) Average indegree outdegree and degree calculation with short equation



For next Monday:

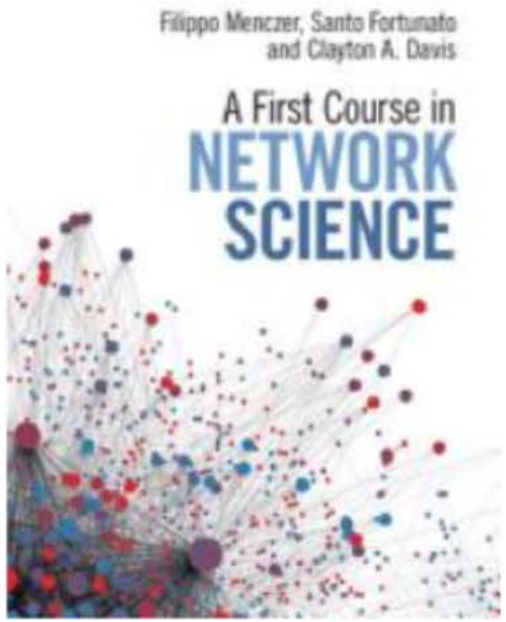
1) Watch Documentary Connected: The Power of Six
   and opt to participate in Lecture 2, see Participation Slides in the Collaboration Module:
   link here:
https://docs.google.com/presentation/d/1iCYxqXHTUKhP1GPSHopvKCQKuLoCRhAxoZbqRktNLPI/edit#slide=id.g92ef02800b_1_7

2) Read the Article that we will cover next Class

3) Finalize the solution of MyFirstNetwork_Exercise-woSoln.ipynb

Textbook:

FIRST COURSE IN NETWORK SCIENCE, ***Author:*** *MENCZER,* ***ISBN:*** *9781108471138*



FIRST COURSE IN NETWORK SCIENCE

Author: MENCZER

ISBN: 9781108471138

Purchase

Rent

Borrow

Purchase

Purchase